Training Evaluation

Overview

Evaluation is a process/procedure that should permeate the entire instructional design process, but we most often think of it as a task to complete at the end of the process or at the end of the delivery of the training. Evaluation is a tool to assess the value and effectiveness of the training program, the results of the training and if the objectives of the training were met. Evaluation is most often accomplished following the delivery of the training giving valuable feedback on the delivery, participant materials, relevance, whether objectives were met, learning that occurred, as well as the effect and impact of the training.

Methods of evaluation may include questionnaires, surveys, interviews, observations, testing (pre/post), and business results analysis. Each methodology can provide effective feedback if designed and executed appropriately.

There are two general categories of evaluation: formative and summative. Formative evaluation refers to the judging of the value of the program while it is in progress and, thus, it focuses on the process. This category of evaluation is most appropriate for evaluating the parts of the program – learning materials, facilitator effectiveness, exercises, etc. – and combining these pieces of data into an overall evaluation.

Summative evaluation focuses on the outcomes of the training program and is usually done at the end of the training or even a time lag after the training. This might include evaluating the business impact of the training program, i.e. cost reduction or revenue generation resulting from the training. As stated below, there are different levels of evaluation. Levels One and Two are examples of formative evaluation, while Levels Three and Four exemplify summative evaluation.

Levels of Training Evaluation

There are several systems for evaluating training programs. One of the most widely used is the Kirkpatrick Model. Using Kirkpatrick's Four Level Evaluation Model you have the ability to gain a full range of feedback useful in validating the training content, delivery, objectives and effectiveness. It also gives the developer the information needed to "fine tune" and revise the training if necessary. As you will see, as the levels of measurement increase so does the complexity of the evaluation methodology as well as the relevance of the feedback.

Level One – REACTION (Formative). This is the most common form of evaluation, and as the name implies, measures how the learners react to the training. This level is most often measured by the use of a questionnaire or survey immediately following the completion of the training. It should be noted that this level of measurement strictly gathers information regarding the participant's perception of the program and not necessarily the effectiveness of the program. However, Level One evaluation is important because it focuses on characteristics of the training presentation that impact participants' interest, attention and motivation, which are critical to the success of any training. Below is a partial example of a Level One evaluation.

Workshop Reaction Survey

Please rate your level of agreement with the following statements using the following scale:

	1= Strongly Disagree	**2** = Disagree	**3**=Agree	**4** =Strongly Agree
1. The workshop met all of its stated objectives.	1	2	3	4
2. Overall, I was satisfied with this workshop.	1	2	3	4

Level Two – LEARNING (Formative). The Learning level refers to the question: Did the participants learn anything? Determining learning may often include testing, both pre- and post-training to isolate what was known prior to the program and what knowledge was gained via the program. This measurement allows us to validate the learning objectives and is a way to assess the participant. In this case the evaluation comes from the analysis of the assessment. As was discussed during Phase 3 of the CDM this level of evaluation is most often developed in that phase and implemented immediately following the completion of the program by the participant.

An example of a Level Two evaluation may be to have a training participant, i.e. a housekeeper, demonstrate the cleaning of a room while the trainer evaluates the performance using a checklist to ensure all key criteria for a clean room were completed satisfactorily. Another example might be to require a participant from a workshop that trained students how to read and evaluate a profit and loss statement to identify the key elements of the profit and loss statement utilizing an online assessment or paper-and-pencil test.

Level Three – IMPACT (Summative). Level three focuses on the evaluation of the behavior change and may be done immediately following completion of the training or when the participant returns to the work environment. This level involves testing participants' capability to perform the learned skills whether in the classroom or on the job. (An ability to perform on the job is going to be a higher level of validation of the program.) Common evaluation applications of this level may be observation—formal (i.e. role play) or informal (i.e. on the job observation)—or through a questionnaire administered after the participant has had ample opportunity to implement the learning on the job. Below is a partial example of a post training follow up survey, which would focus on the training materials' effectiveness in changing behavior on the job.

Post Training Survey

Please rate your level of agreement with the following statements using the following scale:

	1= Strongly Disagree	**2** = Disagree	**3**=Agree	**4** =Strongly Agree
1. I have applied the skills & knowledge from the training in my job	1	2	3	4
2. The skills & knowledge obtained in the training have had a direct effect on my performance	1	2	3	4
…				
…				
…				
10. Please explain or provide examples on how this training has impacted you or your organization's performance.				

_____

Level Four – RESULTS (Summative). This final level of measurement and evaluation focuses on the impact of the training. This level asks what the training has accomplished from a business perspective. Such impacts may include employee retention, morale, revenue increases, expense decreases, enhanced teamwork, etc. It is difficult to isolate the results of a training program in a dynamic business because links cannot be easily made between the training and organizational or individual improvements. As this is the highest level of evaluation, it is the most complex and the most useful. Most business partners want to evaluate the effect of the training on the business, the performance outcome, and possibly the return on the investment (ROI).

For example, imagine a training program was designed to increase the sales/revenue of a hotel. The purpose of Level Four evaluation would be to establish the baseline sales for the individual training participant and then, after a period of time using the same measurement, measure the sales for the individual. The difference between the baseline measurement and the measurement following the training will identify the impact the training had on the business. Note: Training does note take place in a vacuum so the training may only be one of the factors affecting the new measurement.

Using Evaluative Information

Data from each of the levels will provide a thorough analysis of the training program. Consolidating the data of Level One and Level Two can help the developer and facilitator identify the strengths and weaknesses of the training design and delivery. (I.e. which exercises, visuals, etc. were effective, how well did the program keep the participants' attention, how relevant was the content to the participants and their job, etc.) Levels Three and Four can provide some of the same useful information in an even more objective format. Behavioral and Results oriented feedback will more specifically validate the content of the training and whether the content is positively impacting the work environment. This information is of critical importance to many of the stakeholders such as the training sponsors, etc., and can give increased credibility to the training.

Summary

Each level of evaluation provides useful information for the program stakeholders. The first three levels predominantly provide program measurements, and the usefulness of the feedback is predominantly to the training developer and facilitator. They provide significantly for the improvement of the training. The business partners, who most often identify the business need for the training and provide the funding, are often most interested in Level Four, the impact the training had on the business, individual or organization, in other words, the return provided by the training. All levels of measurement are important and a thorough analysis of the evaluation methodology for each training program is necessary before implementation. The purpose of the training, cost and complexity of evaluation methods, etc., will be key factors in determining the methodology used.

Bibliography

Brown, Frederick G (1971). *Measurement and Evaluation*. Itasca, Illinois: F.E. Peacock.

Gilbert, T. (1998). A Leisurely Look at Worthy Performance. *The 1998 ASTD Training and Performance Yearbook*. Woods, J. & Gortada, J. (editors). New York McGraw-Hill.

Hayes, M. (2003, Feb 3). Just Who's Talking ROI? *Information Week*. p. 18.
Kirkpatrick, Donald, (1994). *Evaluating Training Programs*. San Francisco, CA: Berrett-Koehler Publishers, Inc.

Markus, H. & Ruvulo, A. (1990). Possible selves. Personalized representations of goals. *Goal Concepts in Psychology*. Pervin, L. (Editor). Hillsdale, NJ: Lawrence Erlbaum. pp. 211-241.

Tovey, Michael (1997). *Training in Australia*. Sydney: Prentice Hall Australia.